PREDICTION OF THE PCDD/F, dl-PCB, AND TOTAL 2005-WHO-TEQ VALUES ON THE BASIS OF SIX CONGENER CONCENTRATIONS IN FISH: TOWARD A NEW SCREENING STRATEGY FOR THE CONTROL?

Ronan Cariou[1], Philippe Marchand[1], Anaïs Vénisseau[1], Aline Brosseaud[1], Dominique Bertrand[2], El Mostafa Qannari[3], Jean-Philippe Antignac[1] and Bruno Le Bizec[1]

[1]LABoratoire d'Etude des Résidus et Contaminants dans les Aliments (LABERCA), Ecole Nationale Vétérinaire de Nantes (ENVN), USC INRA 2013, BP 50707, 44307 Nantes Cedex 3, France; [2]Unité Biopolymères, Interactions, Assemblages, Institut National de la Recherche Agronomique (INRA), Nantes, France, [3]Unité de Sensométrie et de Chimiométrie, Ecole Nationale d'Ingénieurs des Techniques des Industries Agricoles et Alimentaires (ENITIAA), Nantes, France

Abstract

Current European regulation related to polychlorinated dibenzo-*p*-dioxins/furans (PCDD/F) and dioxin-like polychlorinated biphenyls (dl-PCB) in food is based on WHO-TEQ values. For confirmatory purpose, Gas chromatography coupled with high resolution mass spectrometry (GC-HRMS) according to the isotope-dilution method is usually the method of choice for precisely measuring the 29 target congener in three separated fractions. Time and cost related to these analyses are very significant. Diverse kinds of screening concepts can be envisaged: bioanalytical techniques or still GC-MS related techniques. In the present study, we elaborated and validated a prediction model fit for the GC-MS screening approach in fish, based on the measurement of 4 PCDD/F and 2 co-planar dl-PCB congeners, potentially analyzable in a single extracted fraction. Consequent independent datasets were used for model elaboration (n=108) and validation (n=363, n=357 and n=6), and performances of the model were evaluated in terms of robustness, prediction capability and false negative/positive rates.

Introduction

Current European regulation related to polychlorinated dibenzo-*p*-dioxins/furans (PCDD/F) and dioxin-like polychlorinated biphenyls (dl-PCB) in food is based on WHO-TEQ values, which are linear combination of the concentrations measured for 17 PCDD/F and 12 dl-PCB congeners weighted by their respective Toxic Equivalent Factors[1]. For confirmatory purpose, the precise quantification of these congeners according to the isotope-dilution method is then mandatory[2]. Usually, three purified extracted fractions (each containing PCDD/F, co-planar dl-PCB and mono-ortho dl-PCB, respectively) are prepared from each sample to be analyzed. Gas chromatography coupled with high resolution mass spectrometry (GC-HRMS) is usually the method of choice for measuring the different target congeners in these three fractions, which lead to a significant number of diagnostic signals to interpret. Time and cost related to these analyses are consequently very significant, due to either the sample preparation procedure and data analysis steps[3].

Time and cost saving expected from a new appropriate approach, at least for screening purpose, is also a major issue in European regulation[2]. In this scope, two kinds of screening concepts can be envisaged, *i.e.* bioanalytical techniques including cell-based and kit-based approaches, or still GC-MS related techniques with isotopic-dilution method. In this latter case, two modifications may be proposed: 1- to use less expensive instruments such as GC-MS/MS[5-6], GC-HRTOFMS[6-7] or GCxGC-HRTOFMS[8], 2- to predict the WHO-TEQ values on the basis of a limited number of congeners, preferentially analyzed in a single extracted fraction. A restricted literature is available about regression models and we can only cite a work on environmental matrices[9].

Correlations with non-dioxin-like compounds have also been tentatively proposed, such with indicator PCB in indoor air (n=8)[10], total PCB in fish[11] or fatty acid pattern in fish products[12]. However, to our knowledge, no comparable satisfying model has been developed yet for both PCDD/F-TEQ and dl-PCB-TEQ in food matrices.

In the present study, we elaborated and validated a prediction model fit for the GC-MS screening approach in fish, based on the measurement of 4 PCDD/F and 2 co-planar dl-PCB congeners. The model was applied to independent datasets to evaluate its performances in terms of robustness, prediction capability and false negative/positive rates.

Materials and Methods

Datasets

The elaboration of the predictive model was based on PCDD/F and dl-PCB occurrence data collected in the frame of the 2004 French monitoring plan, including 108 analyzed fish samples from 32 species, excluding *Anguilla anguilla*. The validation of the elaborated model was based on three additional and independent datasets, also excluding *Anguilla anguilla* species. The first one included the results of the 2005 and 2006 French monitoring plans, including 363 analyzed fish samples from 42 species (mostly wild saltwater fishes, but also fishes from aquaculture and freshwater). The second one corresponded to occurrence data generated during two particular studies at the National level in 2007 in the Rhône river area (357 samples, 22 freshwater species), including 3 Alpine lakes (lac du Bourget, lac Léman and lac d'Annecy, respectively with 90, 32 and 28 samples from 7 species). The last one was composed by the consensual values collected from 6 fish samples analyzed through the Interlaboratory Comparisons on Dioxins in Food organized by the Norwegian Institute of Public Health between 2003 and 2007.

Sample analysis

All samples have been analyzed according to the isotope-dilution method with GC-HRMS measurement. Target compounds were the 17 PCDD/F and the 12 dl-PCB. Only filets were analyzed. These analyzes were performed by four different accredited laboratories at French national level.

Statistical analysis

Each PCDD/F and dl-PCB congener was considered as a statistical variable and each analyzed sample as an observation. For a given observation, the value assigned to each variable corresponded to the upper bound concentration measured for the corresponding congener, expressed in $pg.g^{-1}$ fresh weight (f.w.). The 2005 WHO TEQ calculated for PCDD/F, dl-PCB and Total were also considered as variables to be predicted. Since the 2005 WHO TEF are likely to replace 1998 WHO TEF in the next regulation update, it was decided to use it in the present study. Additional informative variables were introduced, including the extracted fat amount and the fish species. Statistical analysis included hierarchical clustering of the variables (using the Ward aggregation method and the $1-\rho$ metric) and step-by-step incremental multiple linear regression, and were realized using Statistica© software (v. 7.1, Statsoft, Inc., Tulsa, USA)..

Results and Discussion

Model elaboration

A step-by-step incremental multiple linear regression was performed on the first dataset (2004 French monitoring plan). Since most of the time PCDDF and dl-PCB are from independent release sources in the environment, these two classes of compounds were considered separately. Results are shown in Table 1.

Concerning PCCD/F, a very good predictive linear model ($R^2 > 0.99$) was already obtained at the third step of the analysis (3 congeners included). The 2,3,7,8-TCCD was then included at step 4, permitting to reach even higher efficiency of the model ($R^2 > 0.9995$). On the basis of these results, it was decided to retain a four congener model for PCDD/F. We can highlight that these four selected congeners are lower chlorinated ones (tetra- and pentachlorinated) and share close physico-chemical properties. Purifying them in a unique pure fraction appears conceivable without major difficulty.

Concerning dl-PCB, the PCB-126 alone led to a very good predictive capability ($R^2 > 0.999$). However, as in previous case of PCDD/F, it was decided to include one more dl-PCB congener in the model to reach a R^2 value higher than 0.9995. Nevertheless, instead of choosing the step 2 congener (PCB-105), it was decided to select the

step 3 congener (PCB-169) which is also a co-planar dl-PCB. This choice can be justified by the aim to make this prediction model compatible with an easy purification protocol. This protocol has to lead to a unique fraction (one injection), pure enough to allow detection of all selected congeners. The possibility to elute co-planar compounds (*i.e.* the 4 tetra- and pentaCDD/F and the 2 co-planar dl-PCB) from a carbon column appears more relevant. Impurities usually eluted with mono-ortho PCB should be removed. For example, another procedure has been proposed for a rapid extraction and purification of PCCD/F and co-planar dl-PCB in a single fraction from serum samples[13].

The final predictive equations were the ones described in Table 2. Considering the hierarchical clustering, the choice of congeners appeared highly coherent. The two chosen dl-PCB congeners were very close to the dl-PCB-TEQ and the 4 chosen PCDD/F congeners appeared to be the closest ones to the PCDD/F-TEQ (data not shown). As shown on the graphical representations of the predicted *versus* observed values (Figure 1), the proposed predictive models were proven to be very satisfactory.

Table 1: Resulting coefficients from the incremental step-by-step multiple linear regressions performed on the 2004 French monitoring plan (*n=108*), for PCDD/F and PCB independently. Each congener concentration was weighted by its 2005-TEF prior to calculations. Grey columns correspond to the finally selected models.

PCDD/F	**2005-TEF**	**Step 1**	**Step 2**	**Step 3**	**Step 4**	**Step 5**	**Step 6**
Constant factor		0.08425	0.04245	0.00213	0.00557	0.00177	0.00051
2,3,7,8-TCDD	1				0.99622	0.98594	0.98863
1,2,3,7,8-PeCDD	1			1.44984	1.17288	1.10197	1.04304
1,2,3,4,7,8-HxCDD	0.1						
1,2,3,6,7,8-HxCDD	0.1						1.42727
1,2,3,7,8,9-HxCDD	0.1						
1,2,3,4,6,7,8-HpCDD	0.01						
OCDD	0.0003						
2,3,7,8-TCDF	0.1		1.41298	1.39330	0.99576	1.02058	1.02099
1,2,3,7,8-PeCDF	0.03						
2,3,4,7,8-PeCDF	0.3	2.37678	1.57172	0.93185	1.00865	0.96257	0.98180
1,2,3,4,7,8-HxCDF	0.1						
1,2,3,6,7,8-HxCDF	0.1					5.50486	3.69723
1,2,3,7,8,9-HxCDF	0.1						
2,3,4,6,7,8-HxCDF	0.1						
1,2,3,4,6,7,8-HpCDF	0.01						
1,2,3,4,7,8,9-HpCDF	0.01						
OCDF	0.0003						
R^2		0.88596	0.97804	0.99213	0.99972	0.99992	0.99998
dl-PCB	**2005 TEF**	**Step 1**	**Step 2**	**Step 3**	**Step 4**	**Step 5**	**co-planar only**
Constant factor		0.01458	0.00806	0.00275	0.00221	0.00052	0.00527
PCB 77	0.0001				0.98120	1.02684	
PCB 81	0.0003						
PCB 126	0.1	1.11147	1.03138	1.02266	1.01992	0.99911	1.08022
PCB 169	0.03			0.63622	0.80732	1.09681	0.93607
PCB 105	0.00003		6.58464	5.55506	4.87512	0.34906	
PCB 114	0.00003						
PCB 118	0.00003					1.48823	
PCB 123	0.00003						
PCB 156	0.00003						
PCB 157	0.00003						
PCB 167	0.00003						
PCB 189	0.00003						
R^2		0.99918	0.99970	0.99984	0.99993	0.99998	0.99950

Table 2: Final predictive equations selected on the basis of 4 PCDD/F and 2 PCB congeners. Each TEQ, expressed in pg.g^{-1} f.w., is the sum of the obtained factor (left) multiplied by the congener concentration expressed in pg.g^{-1} f.w. multiplied by its corresponding 2005-TEF.

Predicted PCDD/F-TEQ					
=	0.00557				
+	0.99622	*	1	*	[2,3,7,8-TCDD]
+	1.17288	*	1	*	[1,2,3,7,8-PeCDD]
+	0.99576	*	0.1	*	[2,3,7,8-TCDF]
+	1.00865	*	0.3	*	[2,3,4,7,8-PeCDF]
Predicted dl-PCB-TEQ					
=	0.00527				
+	1.08022	*	0.1	*	[PCB-126]
+	0.93607	*	0.03	*	[PCB-169]
Predicted Total-TEQ					
=	0.01084				
+	0.99622	*	1	*	[2,3,7,8-TCDD]
+	1.17288	*	1	*	[1,2,3,7,8-PeCDD]
+	0.99576	*	0.1	*	[2,3,7,8-TCDF]
+	1.00865	*	0.3	*	[2,3,4,7,8-PeCDF]
+	1.08022	*	0.1	*	[PCB-126]
+	0.93607	*	0.03	*	[PCB-169]

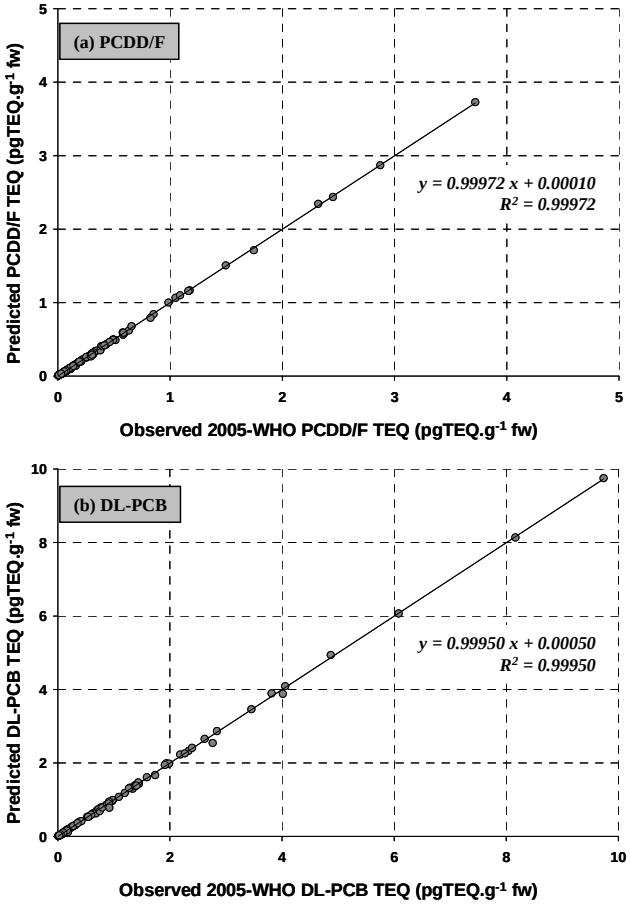

Figure 2: Correlation between observed and predicted PCDD/F-TEQ (a) and dl-PCB-TEQ (b) values obtained for the 2004 French monitoring plan (n=108 fish samples) on the basis of the model elaborated by the multiple linear regression.

Model validation

In order to validate the proposed predictive model, the obtained equations were applied to the three independent datasets described above. In all cases, the results demonstrated an excellent correlation between the predicted and observed 2005-PCDD/F-TEQ and dl-PCB-TEQ values ($R^2 > 0.99$). Only 2 atypical observations were excluded for PCB regressions, these two samples presenting unusual higher proportion of PCB-118. In order to precise the eventual error induced by this predictive model, all the deviations between predicted and observed values were calculated. Figure 2 presents the results obtained for the three validation datasets.

Globally, the deviations distribution highly widens when observed values decreased. Since the standard deviation related to the confirmatory method is not directly dependent on the TEQ level, a lower denominator (observed value) logically results in a higher deviation. Considering that these higher deviations did not result in false positives, this bias can be considered as of no significance for the regulatory purpose. Assuming that maximum levels allowed by European Union for fish are set at 4 and 8 pg TEQ.g^{-1} f.w. for PCDD/F-TEQ and Total-TEQ (respectively), observed limits of 0.75 pg TEQ.g^{-1} f.w. for PCDD/F-TEQ and 1.5 pg TEQ.g^{-1} f.w. for dl-PCB-TEQ were considered regarding the interpretation of the predicted results. Above the considered limit, all the deviations between predicted and observed PCDD/F-TEQ values were found to be lower than 12% (one exception at 15%). Concerning dl-PCB-TEQ, 28 samples led to deviations between predicted and observed values higher than 12%, but none above 21%. 27 of these 28 samples were from the Rhône river area. This might be explained by the fact that samples included in the

elaborated model were not mainly river fish. A regression curve for dl-PCB based on the Rhône river area dataset revealed that PCB-118 moved from the fifth to the second place as important congener, improving deviation dispersion. Such importance of PCB-118 congener was not observed in river fish from French monitoring plans. In a final validation step, the samples analyzed within the framework of interlaboratory comparisons (third dataset) demonstrated once again a very good efficiency of the proposed predictive models. Deviations were lower than 12% for PCDD/F-TEQ and lower than 4% for dl-PCB-TEQ and Total-TEQ.
Comparatively, results obtained on the basis of 1998-TEF were found less accurate, especially concerning dl-PCB (data not shown). This observation can be linked to a lower relative weight of co-planar dl-PCB in dl-PCB-TEQ, a mathematical phenomenon already observed[14-15]. Introduction of the 2005-TEF in the next European Union regulation update should promote the emergence of regression models fit for the screening purpose.

Application for screening purpose
Maximum levels in fish meat (excluding eel) set by the European Regulation 199/2006 are set at 4 and 8 pg TEQ.g^{-1} f.w., for PCDD/F and sum PCDD/F+dl-PCB, respectively. Considering an action level set at 2/3 the maximum level as values to be screened by the predictive model (*i.e.* respectively 2.67 and 5.33 pg TEQ.g^{-1} f.w.), we fixed a screening cut-off level at 25% under these action levels (*i.e.* respectively 2 and 4 pg TEQ.g^{-1} f.w.), in order to detect samples above the action level. The application of this criterion on the dataset (363 samples from 2005 and 2006 French monitoring plans) appeared highly efficient. Indeed, 5% of these samples were declared as suspect, including only 50% false positives. No false negative sample appeared.

Acknowledgments

The authors want to express their special thanks to the General Directorate for Food (Direction Générale de l'Alimentation, DGAl), to the International Commission for the Protection of Lake Geneva (CIPEL) and to the Syndicat Mixte du Lac d'Annecy (SILA) for providing and allowing us to exploit the datasets.

References

1. Commission Regulation (EC) No 1883/2006 of 19 December 2006.
2. Commission Regulation (EC) No 199/2006 of 3 February 2006.
3. Reiner E.J., Clement R.E., Okey A.B. and Marvin C.H. *Anal Bioanal Chem* 2006;386:791-806.
4. Aristizabal B., Quintero A., Cobo M., Suarez N., Hoyos E., Montes de Correa C., Avalos M., Abad E. and Rivera J. *Organohalogen Compounds* 2007;69:1269-1271.
5. Dobrzynska E. and Szewczynska M. *Organohalogen Compounds* 2007;69:1333-1336.
6. Tagasuga T., Takemoni H., Matsukami H. and Inoue T. *Organohalogen Compounds* 2007;69:1134-1137.
7. Matsukami H., Takemoni H., Takahashi S., Tagasuka T. and Tanabe S. *Organohalogen Compounds* 2008;70:46.
8. Hashimoto S., Takazawa Y., Fushimi A., Ito H., Tanabe K., Noma Y., Shibata Y., Ubukata M., Kusai A. and Tanaka K. *Organohalogen Compounds* 2007;69:1106-1109.
9. Hayashi S., Goto S., Hirano M., Ohtsuka K. and Furuno N. *Organohalogen Compounds* 2007;67:1002-1005.
10. Heinzow B., Mohr S., Ostendorp G., Kest M. and Körner W. *Chemosphere* 2007;67:1746-1753.
11. Bhavsar S.P., Hayton A., Reiner E.J. and Jackson D.A. *Environ Toxicol Chem* 2007;26(8):1622-1628.
12 Bassompierre M., Tomasi G., Munck L., Bro R. and Engelsen S.B. *Chemosphere* 2007;67:S28-S35.
13. Focant J.F., Eppe G., Massart A.-C., Scholl G., Pirard C. and De Pauw E. *J Chrom A* 2006;1130:97-107.
14. Wittsiepe J., Fürst P. and Wilhelm M. *Int J Hygiene Environ Health* 2007;210:335-339.
15. Hasegawa J., Guruge K.S., Seike N., Shirai Y., Yamata T., Nakamura M., Handa H., Yamanaka N. and Miyazaki S. *Chemosphere* 2007;69:1188-1194.

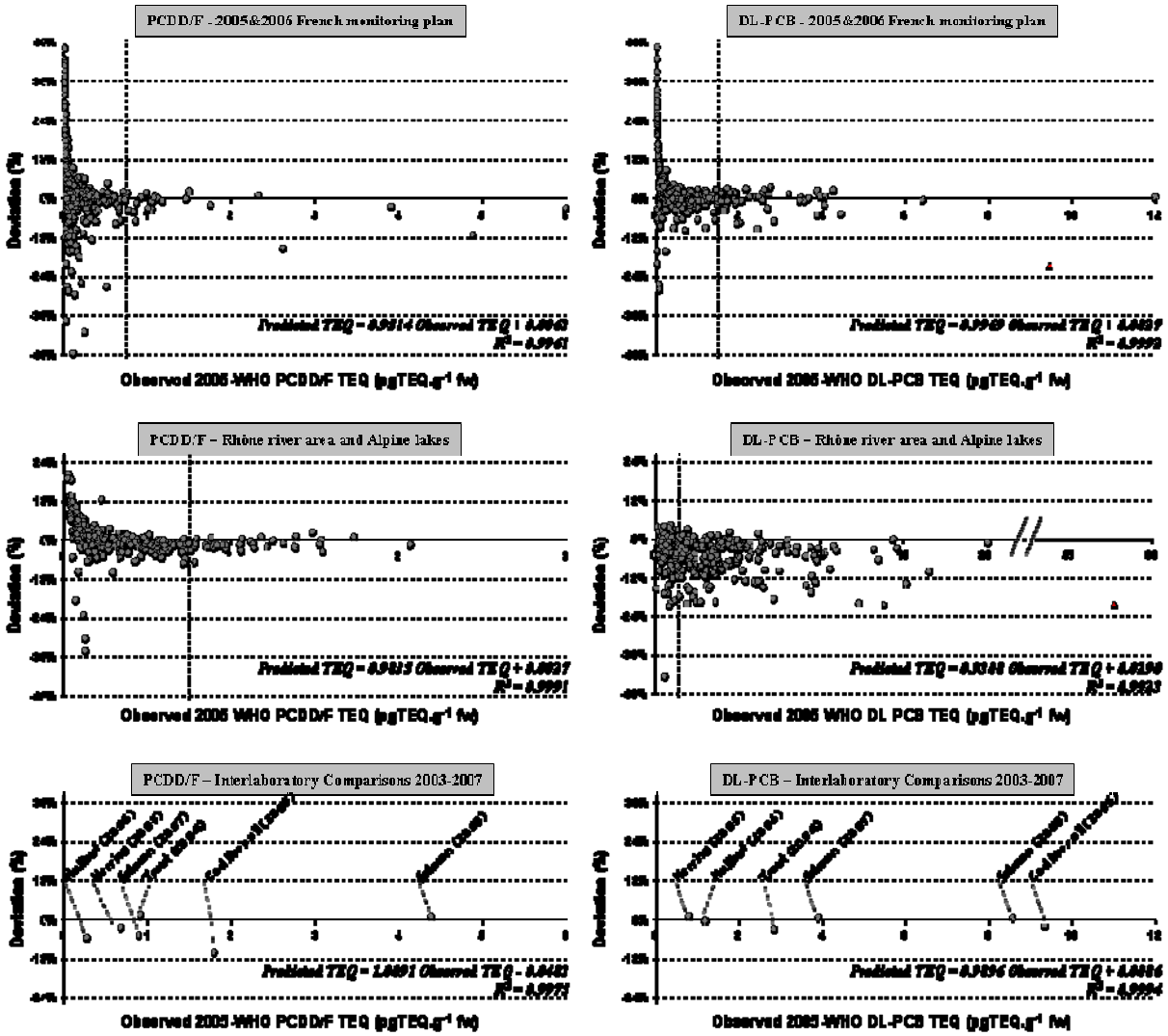

Figure 2: Deviations calculated between predicted and observed PCDD/F-TEQ (left) or dl-PCB-TEQ (right) for the 2005-2006 French monitoring plan dataset (up, n=363), the Rhône river area and Alpine lakes dataset (middle, n=357) and the consensual data of the Interlaboratory Comparisons on Dioxins in Food with sample identification (down, n=6). Corresponding regression curve equations between predicted and observed TEQ are also specified for each figure. Dashed vertical lines: set interpretation limits. Red triangles: the 2 atypical samples excluded for PCB regressions.